CHAPTER 3
LINK-BUDGET AND SYSTEM SPECIFICATIONS

In this chapter, link-budget design equations and specifications of the lidar system are derived and discussed.

Setting off from the basic lidar equation, emphasis is given to the system optics trade-offs, basically, emitting laser and receiving telescope. Here, the design is further complicated by geometrical considerations that include the degree of overlap between the laser beam and the field-of-view of the receiver optics. Typical power levels, signal-to-noise ratio at the receiver's output and lidar range performance are also assessed. This involves evaluation of different photodiodes and system operation modes such as shot-noise limited, dark-noise limited and thermal-noise limited.

The interpretation and cross-examination of the great quantity of lidar parameters that play a role in the overall system performance has justified the development of a second software package, *link-budget*. The plots and computations presented have been worked out with it and include measured values whenever possible. In addition to the design equations, there are economical, technological and mechanical constraints that limit the scope of the equations derived. Other chapters of this thesis work will concentrate on the design and measurement steps of the devices involved.

1. LINK-BUDGET ORIENTED LIDAR EQUATIONS

1.1 The basic lidar equation

Since this project is concerned to the elastic interaction between the laser radiation and the atmosphere, where both incident and scattered radiation are at the same wavelength, the general form of the lidar equation [25] may formally be expressed in the following *single-scattering lidar equation* [9]:

$$P(\lambda,R) = P_o\left(\frac{c\tau_l}{2}\right)\beta(\lambda,R)\frac{A_r}{R^2}\exp\left[-2\int_0^R \alpha(\lambda,r)\,dr\right]\xi(\lambda)\,\xi(R) \tag{1}$$

where:

$P(\lambda,R)$ is the instantaneous received power from the range R and wavelength λ,

P_o is the transmitted power at time t_o,

c is the speed of light,

τ_l is the laser pulse duration,

$\beta(\lambda,R)$ is the volume backscatter-coefficient of the atmosphere [$m^{-1}\cdot sr^{-1}$] at wavelength λ over range R,

R is the range,

A_r is the effective receiver area,

$\alpha(\lambda,R)$ is the volume extinction-coefficient of the atmosphere [m^{-1}] at wavelength λ over range R,

$\xi(\lambda)$ is the receiver's spectral transmission factor and,

$\xi(R)$ is the overlap factor or geometrical form factor.

Before proceeding further, it is worth making some comments about the interpretation of the coefficients α and β in the lidar equation:

The volume extinction coefficient, α, is defined as the cross-section per unit volume $[m^2/m^3]$, and it is numerically identical to the linear extinction-coefficient used in laser propagation equations, which has units of $[m^{-1}]$. It also coincides with the extinction-coefficient introduced in Chap.2.

The volume backscatter-coefficient, β, is defined as the fractional amount of incident energy scattered per unit solid angle [sr] in the backward direction per unit atmospheric length. It agrees with the definition given in Chap.2 as the backscatter cross-section per unit volume and unit solid angle $[m^2/(m^3 \cdot sr)]$.

Since the transmitted laser pulse has a finite duration τ_1, it illuminates a finite geometrical length $c\tau_l$ of the atmosphere at any instant. However, since the received energy must travel a two-way path, the atmospheric length (or range increment) from which signals are received at any time t is just half this value. Thus, the distance

$$\Delta R = \frac{c\tau_l}{2} \tag{2}$$

is called the *effective pulse length*. It is exactly half of the actual length instantaneously illuminated by the pulse because of the two-way path that all received energy must traverse.

There are, however, other parameters of concern to the receiver and emitting optics. The effective receiver area A_r enters eq.(1) because it determines the solid angle

$$\Omega_o = \frac{A_r}{R^2} \qquad [sr] \tag{3}$$

subtended by the receiver at range R [26]. *Provided that the angle of divergence of the transmitter beam does not exceed the angle of acceptance of the receiver, all the energy transmitted will be available for backscattering at range R.* Otherwise *overlap factor losses should be considered.*

The expression

$$T(\lambda,R) = \exp\left[-2\int_0^R \alpha(\lambda,r)\,dr\right] \tag{4}$$

represents the fractional transmittance of energy at wavelength λ along the two-way path between the lidar and range R. This depends upon α, the volume extinction coefficient of the atmosphere, which represents the fraction by which the flux of energy in the direction of propagation is reduced per unit length. From Chap.2, gases, aerosols and hydrometeors contribute to both absorption and scattering of energy. Formally,

$$\alpha = \alpha_{g,scat} + \alpha_{g,abs} + \alpha_{p,scat} + \alpha_{p,abs} \tag{5}$$

where the subscript g means gases (Rayleigh's components, see Sect.2, Chap.2), p means particulate constituents (Mie's components, see Sect.3, Chap.2), *abs* means absorption and *scat* means scattered. Recall that $\alpha_{g,a}=0$ at the operating wavelengths of the lidar, $\lambda=532nm$ *and* $1064nm$ (Sect.1, Chap.2). Strictly speaking, the term α applies in this simple form only when Bouguer's law (the extinction law) is valid, that is, when essentially all the scattered energy is permanently removed from the outgoing lidar beam. In strongly scattering atmospheres (e.g. clouds, thick hazes, fog) multiple scattering occurs, and corrections must be made to eq.(1).

Alternatively, the lidar equation may be formulated in terms of the emitted energy noting that [25]

$$P(\lambda,R) = \frac{E(\lambda,R)}{\tau_d} \quad ; \quad P_o = \frac{E_l}{\tau_l} \tag{6}$$

where

$P(\lambda,R)$ is the range-received power at wavelength λ,

$E(\lambda,R)$ is the range-received energy at wavelength λ,

P_o is the laser transmitted power at time t_o,

E_l is the laser pulse energy,

τ_l is the laser pulse duration and,

τ_d is the sampling time or integration period of the detector.

It becomes:

$$E(\lambda,R) = E_l \frac{c\tau_d}{2}\beta(\lambda,R)\frac{A_r}{R^2}T(\lambda,R)\xi(\lambda)\,\xi(R) \tag{7}$$

Finally, the terms $\xi(\lambda), \xi(R)$ represent the receiver's optical losses. Thus, the first one, represents the receiver's spectral transmission including any spectrally selecting elements such as filters or monochromators. Because of the high degree of monochromaticity of the laser energy, extraneous light can be excluded to a substantial degree by the use of a narrow band filter centered on the laser wavelength. Basically, such light will come from the background radiation (some figures were estimated at the operating wavelengths by *link-atmos* in Sect.5, Chap.2 and are reproduced here in Tab.1). The increment of radiative background power accepted by the receiving optics can be expressed in the form [26]

$$P_{back} = L_b(\lambda)A_r\Omega_r d\lambda \qquad [W] \tag{8}$$

where

$L_b(\lambda)$ represents the spectral radiance of the sky background [$W \cdot m^{-2} \cdot \mu m \cdot sr^{-1}$],

A_r is the receiver-system effective aperture [m^2]

Ω_r is receiver-system acceptance solid angle (eq.(3)) [sr]

$d\lambda$ is the equivalent spectral width of the spectrally selecting component (e.g. filter) [μm].

Note that the term P_{back} represents the incident background power within the optical bandwidth and hence the receiver-system transmission efficiency $(\xi(\lambda),\ \xi(R))$ is computed separately (eq.(1)). In other words, P_{back} is only affected by the factor $\xi(\lambda)$.

The second term or *overlap factor*, $\xi(R)$, depends only upon the overlap of the area of laser irradiation (spot at range R) with the field of view of the receiver optics. Incidentally, because laser energy can be highly collimated, it would be desirable to direct all the transmitted energy into a narrow beam and restrict the field of view of the telescope to this angle, thus minimizing background light entering the system. The measured divergence is about *0.1 mrad* for the Nd:YAG laser used but typical values may also reach 2 or 3 mrad. In the following chapters, this point will be reconsidered since the election of our photodiode may severely limit the field-of-view.

1.2 Simple overlap factor

The geometrical form factor for a coaxial lidar having no apertures (other than the objective lens or mirror of the telescope) or obstructions is unity, provided the divergence angle of the laser beam is less than the opening angle of the telescope.

The geometrical form factor can be evaluated fairly easily for a biaxial lidar (Fig.1), as the present case is, if the objective lens of the telescope represents the limiting aperture of the receiver optics, any obstruction is neglected, and an uniform laser distribution over the area of illumination is assumed.
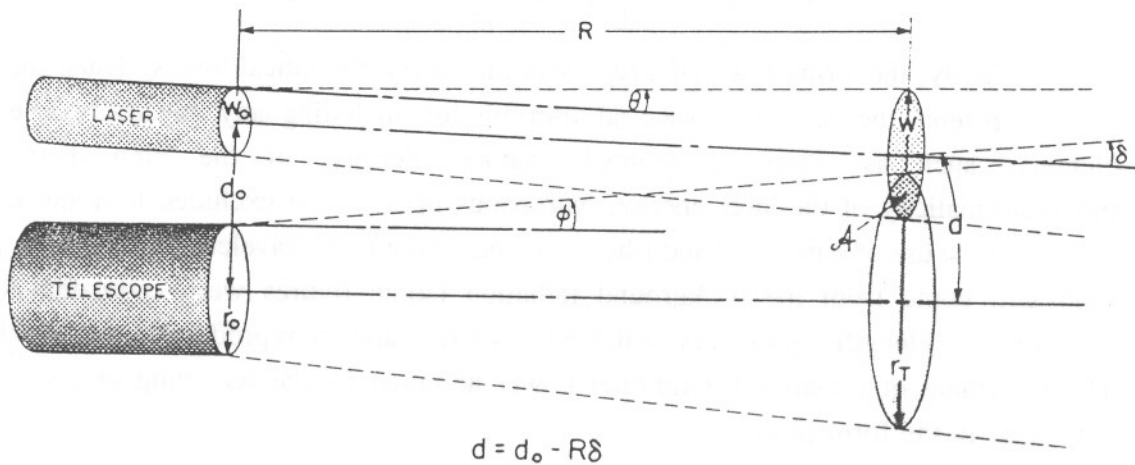

$$d = d_o - R\delta$$

Fig.1 Geometry of the biaxial lidar of the UPC [25].

Under the conditions of Fig.1 one can think of $\xi(R)$ as a simple overlap factor, and one can write [25]

$$\xi(R) = \frac{A\{r_T(R), W(R); d(R)\}}{\pi W^2(R)} \tag{9}$$

where A represents the *area overlap function*, $r_T(R)$, the radius of the radius of the receiver-optics field of view in the target plane at range R, and $W(R)$ the radius of the circular region of laser illumination or spot at range R. In formulating these equations, the field of view of the receiver optics in the target plane is assumed a circle of radius

$$r_T(R) = r_o + \phi R \qquad (10)$$

where

r_o represents the effective radius of the telescope lens (or mirror); and

ϕ represents the receiver-optics half opening angle.

As for the laser beam radius, assuming the output of the laser is predominantly in the TEM_{00} mode, one can write [33][36]

$$W(R) = \{W_o^2 + \theta^2 R^2\}^{1/2} \qquad (11)$$

where

W_o represents the laser output aperture radius; and

θ represents the laser's half divergence angle.

The separation of the telescope and laser axes in the target plane is

$$d = |d_o - R\delta| \qquad (12)$$

where

d_o represents the separation of the axes at the lidar, and

δ represents the inclination angle between the laser and the telescope axes (see Fig.1).

These variables are indicated in Fig.1 along with the laser divergence half angle θ and the telescope half angle field-of-view ϕ. Thus, if a circular detector of radius r_D is positioned on the axis of the telescope of effective focal length f, then [26]

$$\phi = \frac{r_D}{f} \qquad (13)$$
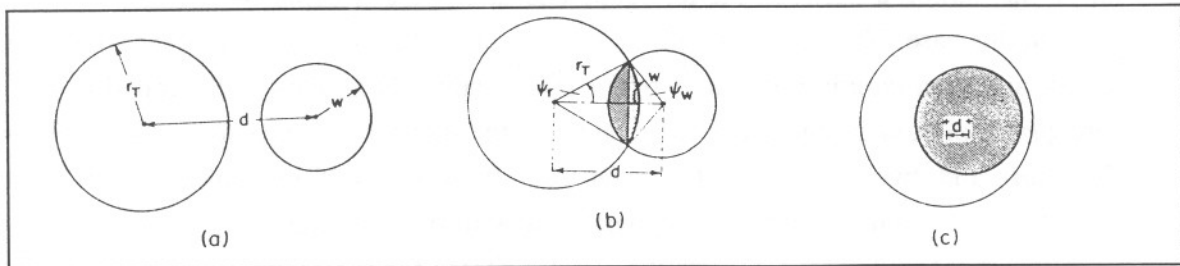
Fig.2 Three overlap situations possible for a biaxial lidar [25].

With regard to Fig.2, three situations are possible:

1. The separation of the axes is too large for there to be any overlap area between the receiver-optics field-of-view and the area of laser illumination (Fig.2a). Then,

$$A = 0 \quad if \quad d > r_T + W \qquad (14)$$

2. *The separation of the axes is small enough that either the area of laser illumination lies totally within the receiver-optics field of view or vice versa.* The former case is illustrated in (Fig.2c) and would correspond to an overlap factor equal to unity. This amounts to saying that if

$$d < |r_T - W| \tag{15}$$

then the overlap factor may be computed as the ratio of areas

$$\xi(R) = \frac{\pi r_T^2(R)}{\pi W^2(R)} \approx \left(\frac{\phi}{\theta}\right)^2 \tag{16}$$

3. *The separation of the axes lies between these extremes (Fig.2b):*

$$|r_T - W| < d < r_T + W \tag{17}$$

Under these circumstances the are overlap function is

$$A\{r_T, W; d\} = W^2 \psi_W + r_T^2 \psi_r - r_T d \sin \psi_r \tag{18}$$

where

$$\psi_W = \cos^{-1} \left[\frac{d^2 + W^2 - r_T^2}{2Wd} \right] \qquad \psi_r = \cos^{-1} \left[\frac{d^2 + r_T^2 - W^2}{2r_T d} \right] \tag{19}$$

It is customary to define the normalized range z, and other dimensionless parameters as follows:

$$z \equiv \frac{R}{r_o}, \qquad A = \frac{r_o}{W_o}, \qquad D = \frac{d_o}{r_o} \tag{20}$$

Based on the normalized range z, the behaviour of $\xi(z)$ in the case of a biaxial lidar with an inclination δ between the telescope and laser axes is depicted in Fig.3.

In this case $\xi(z)$ increases with z at first, attains the value of unity over some range interval, and then rapidly drops to zero for larger values of z. This is illustrated in Fig.3, where $\xi(z)$ is plotted against z for three values of the inclination angle *($\delta=0.05$, 0.01, and 0.001 rad)* and two values of A (5 and 20). In this set of examples, $D=1.25$ and $\theta=\phi=1mrad$, which correspond to typical values in lidar systems.

Besides Fig.3, Fig.4 depicts the overlap factor for the inclinations *$\delta=0$, 1 and 2 mrad* and the telescope and laser defaults of Tab.1. (Range is not normalized).

These results clearly indicate that *interpretation of short-range lidar measurements must take proper account of the overlap factor to compensate the range-received power.* Otherwise, $\xi(z)$ may lead to misinterpretation of the measurements by the inversion algorithms. *Misalignment of the telescope and laser axes is extremely critical.* Thus, simulation using the same default values for *D, θ and φ* have shown that misalignment of only 10 mrad would lead to a lack of signal return for *$z > 300$ (i.e., $R > 300r_o$).*

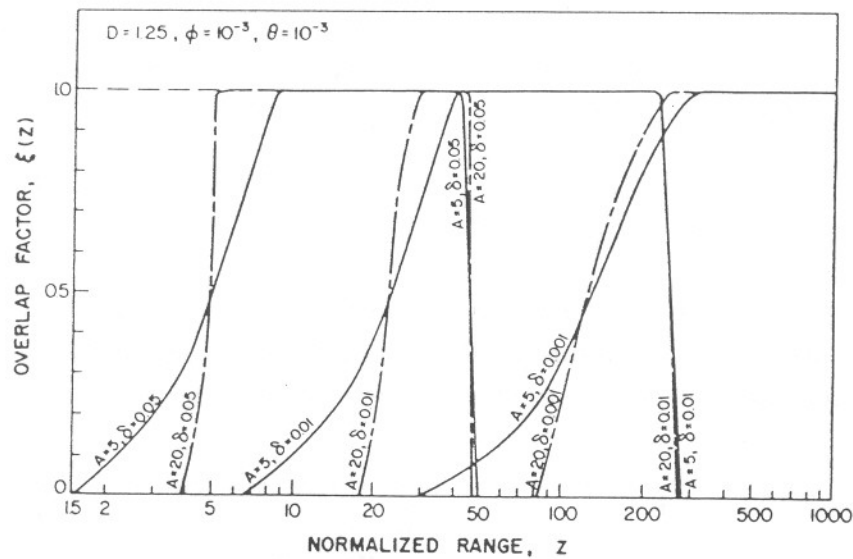*Fig.3 Overlap factor ξ(z) with z (A=5 and 20, δ=0.001, 0.01, and 0.05 rad) [25].*

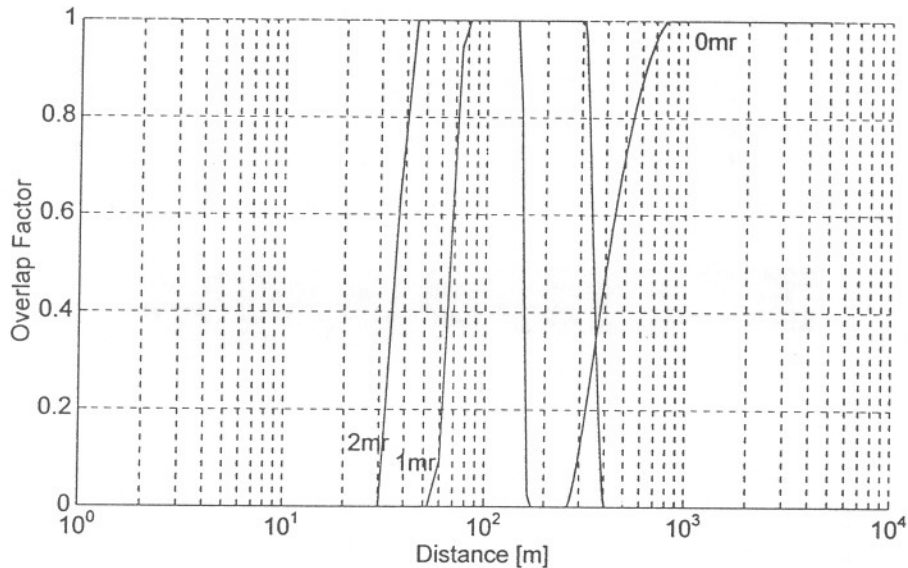*Fig.4 Overlap factor ξ(R) with R (d_o=18 cm and δ=0, 1 and 2 mrad) (Tab.1).*

1.3 The visibility margin approach

Two factors in eq.(1), the elastic volume backscatter-coefficient β, and the attenuation (or extinction) coefficient α determine the lidar signal strength.

The visibility margin concept along with Fig.5 from [9] have been used to obtain lidar performance estimates under various meteorological conditions. Homogeneous atmosphere is assumed [9][25] . The figure shows typical values of β and α for a range of wavelengths and atmospheric conditions. These values have been corroborated using Mie's models and the procedures introduced in Chap.2.

Strictly speaking, the *visibility margin* term is defined under nonturbid atmospheric conditions. That is, one which is free of aerosol particles. Then, the total attenuation of laser radiation is defined by its molecular extinction. The Koshmieder's relation connects

the meteorological visibility V_M with the extinction-coefficient. It is usually given for the wavelength of $0.55\mu m$ (maximum sensitivity of the human eye [9]). Kruse et al. [25] gave a more general relation for any wavelength λ

$$\overline{\alpha_\lambda} = \frac{3.912}{V_M}\left(\frac{550}{\lambda}\right)^q \ [km^{-1}] \qquad q = \begin{cases} 0.585V_M^{1/3} & V_M \le 6km \\ \\ 1.3 & otherwise \end{cases} \qquad (21)$$

The case $q=1.3$ represents average seeing conditions. Expected mean extinction, backscatter and visibility margins are shown decadically spaced in Fig.5, where eq.(21) has even been applied to quite turbid atmospheres such as dense water clouds.
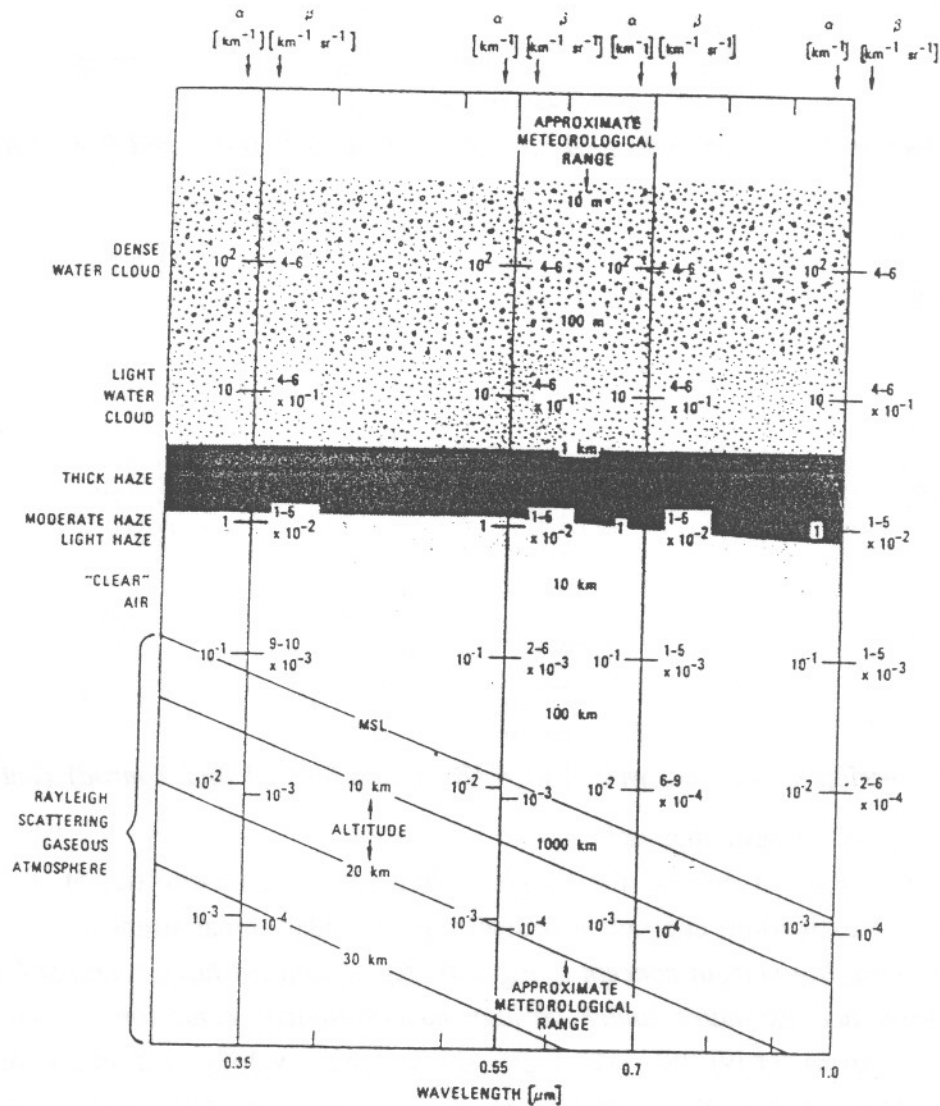

Fig.5 Estimation of extinction, backscatter and visibility margin
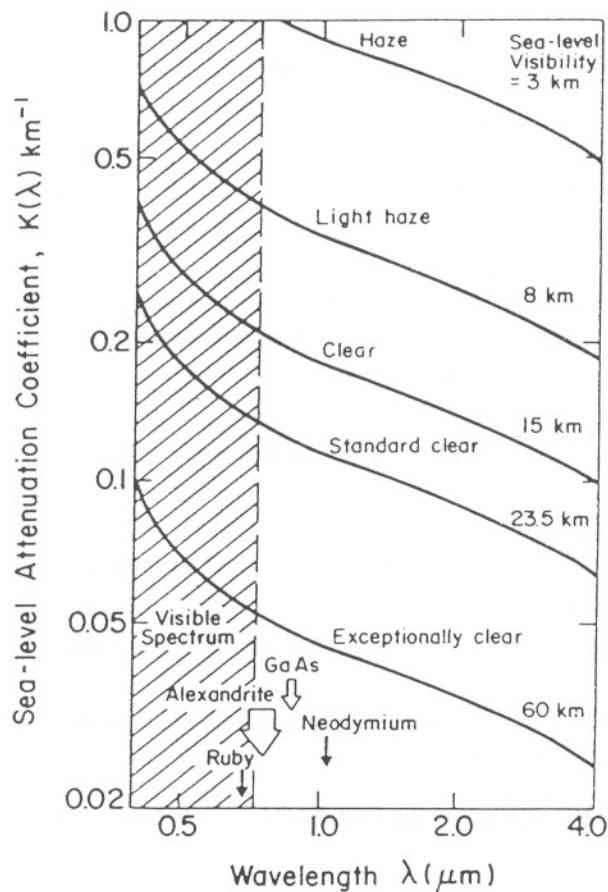variation with atmospheric condition used in link-budget studies [9].

Fig.6 Sea-level attenuation coefficient α(λ) vs. visibility [25].

Fig.6 shows the same concept from another source (Pressley, [25]). Note that more optimistic values are shown.

1.4 Design comments and default values concerning this section

Fig.7 and Fig.8 illustrate range-received power for the visibility margins indicated (values are logspaced) according to the defaults of Tab.1. It includes two situations: in one case, the laser output energy is *500 mJ (λ=532 nm)* and in the other *1 J (λ=1064 nm)*.

From the beginning, a design goal was to choose a photodiode that could either work as PIN and as avalanche photodiode (APD). In addition, it should have acceptable responsivities at both operating wavelengths of the lidar system (1064 and 532 nm). The selection steps led to the EG&G C30954E photodiode.

This criterion has justified the sacrifice of *overlap-factor* considerations in favour of versatility. Another important drawback is the difficult manoeuvrability of the laser. This has conditioned different set-ups for the lidar system that will be discussed in detail along the experimental chapters of this work. For the time being, overlap factor losses have been neglected in Fig.7 and Fig.8 and assume $\xi(\infty)=1$ as shown in Tab.1.

OPTICAL PARAMETERS

Laser parameters: [118]

Energy: $E_{1,1}=500mJ$ ($\lambda_1=532$ nm) $E_{1,2}=1J$ ($\lambda_2=1064$ nm)

Divergence: $\theta=0.1$ mrad   Pulse length: $\tau_1=7\text{-}10$ ns   Waist: $W_o=1$ cm

Receiving optics:

a) Telescope: focal length: $f=2$ m   aperture: $\varnothing20$ cm

                                transmittance: 60 % (includes photodiode window)

                                field-of-view ϕ(eq.(13)): $\phi=0.2$ mrad (photodiode $\varnothing0.8$ mm)

b) Filters [117]:   bandwidth: $\Delta\lambda=10$ nm

                                transmittance: 50 % (at both wavelengths)

Lidar equation parameters (eq.(7))

   Net losses: $\xi(\lambda)=0.3$ (at both wavelengths)

        Overlap factor (eq.(16)): $\xi(\infty)=1$ (if $|\delta|<\theta+\phi$)

        Estimated background radiance: $L_{back}(\lambda_1)=48.86$ W·m^{-2}·μm^{-1}·sr^{-1}

        (Worst-case *global* radiance assessment with *link-atmos* modelling software.

   April 15th, 15:00 GMT, telescope's elevation 90°)

Tab.1 Defaults used in link-budget simulations of Sect.1

As for the background power, there has been assumed the one that will be collected by a vertically aimed telescope under the worst day and hour in April.

Fig.7 and Fig.8 are parameterized for the set of visibility margins *($V_M = 1, 3, 5, 7$ and 10 km). The design has been worked out with these meteorological ranges.*
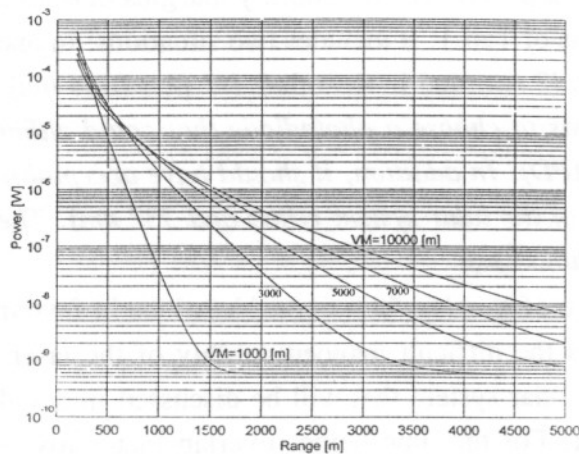
Fig.7 Range-received power ($E_1=500mJ$)
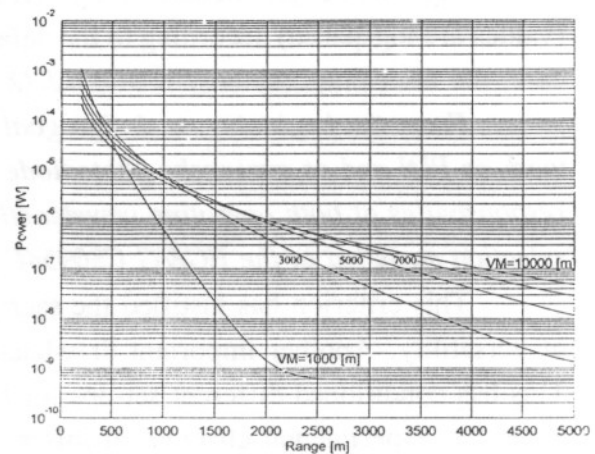
Fig.8 Range-received power ($E_2=1J$).

2. SIGNAL-TO-NOISE RATIO ASSESSMENT IN THE LIDAR RECEIVER

An APD/amplifier combination will be used to detect the faint return of the scattered lidar pulse. Receiver bandwidth, *B,* is typically designed so that $B \approx 0.35/\tau_d$ [149]. Yet, for optimal timing resolution of the return pulse, higher bandwidths as large as $B \approx 0.5/\tau_d$ according to the Nyquist criterion are used. SNR depends on the collecting optics and receiver combination. For optimum SNR, a design/cost compromise must be struck between good optics (low f-number), detector diameter and preamplifier type.

This section focus on the receiver's SNR itself. On the one hand, noise sources in the photodetector along with its intrinsic parameters are reviewed in Sect.2.1. The *transimpedance amplifier (TIA)* will carry out the current-to-voltage conversion of the signal delivered by the photodiode. After that, several stages of amplification will bring the signal amplitude to the level required for further signal processing (such as inversion of the optical parameters, *extinction and backscatter*) in a computer. In order to characterize the lidar receiver (TIA stage in the simplest case), noise sources due to subsequent amplification and conditioning stages will be considered in terms of their equivalent thermal noise variances.

On the other hand, signal-to-noise ratio of the whole lidar receiver will be evaluated in Sect.2.2.

2.1 Noise statistics for an APD-amplifier module

The probability of photoelectric conversion of an incident photon; (or *quantum efficiency η*) in the APD is a function of its wavelength. About *1100 nm,* photons do not possess enough energy to promote silicon valence electrons across the bandgap to the conduction band. Below *1100 nm,* photons of different wavelengths will be absorbed at various mean depths in the APD, depending on the semiconductor's absorption coefficient. Reference [153] provides an excellent insight into semiconductor physics. To improve efficiency, antireflective coatings are evaporated on the diode front surface. By selecting the optical thickness of this layer, quantum efficiency curves are enhanced in the spectral regions of interest [136]. The decisive features of APDs as compared to photomultiplier tubes (PMTs) are reduced price and higher quantum efficiency over a broader spectrum that any PMT [136][122]. For APDs, high gain in linear mode comes from applying a high reverse bias voltage which in turn creates a wide avalanche region around the junction of the semiconductor. For PIN photodiodes, there is not such multiplication [145][147].

The *current (intensity) responsivity* R_i, quantifies the photoelectric gain of a photodetector. It directly depends on the quantum efficiency and it is expressed as

$$R_i = \eta \frac{q\lambda}{hc} M = \frac{\eta \lambda}{1.24 \cdot 10^{-6}} M \qquad \left[\frac{A}{W}\right] \qquad (22)$$

where

> c is the speed of light,
> q is the electron charge,
> λ is the wavelength of the incident photons,
> h is the Planck's constant,
> M is the multiplication factor. ($M=1$ for PIN photodiodes).

Very often, the current responsivity is expressed terms of the intrinsic responsivity of the APD. It is the one the APD would have if it had no gain (PIN mode, $M=1$)

$$R_{io} = \eta \frac{q\lambda}{hc} \quad \rightarrow \quad R_i = R_{io}M \tag{23}$$

If the average number of incident photons on the APD for a time τ_d is k, the average number of photons (after multiplication) is $\mu = M\eta k$. Photoelectron conversions are ruled by Poisson statistics and hence, mean and variance coincide. Yet, as this variance derives from the random fluctuations in the number of carriers generated, the excess noise factor, F, takes into account the randomness of the multiplying phenomenon giving $\sigma^2 = FM^2\eta k$ [123].

To these electrons, there also add those giving rise to the dark current, whose statistics are also poissonian and whose mean and variance are given by

$$\begin{cases} \mu = (I_{ds}+MI_{db})\dfrac{\tau_d}{q} \\[2mm] \sigma^2 = (I_{ds}+FM^2I_{db})\dfrac{\tau_d}{q} \end{cases} \tag{24}$$

where

> I_{ds} is the APD's surface dark current,
> I_{db} is the APD's bulk dark current and the other terms have already been defined.

P-I-N or other types of junctions of a photodiode does not present an infinite resistance to reverse current flow. Consequently, when a reverse bias is applied to a photodiode, a small current flows even in the absence of incident light. This dark current I_d, increases slowly with increasing reverse bias. For an APD, bulk leakage current, I_{db}, is multiplied by the gain, M. Therefore, dark current I_d equals

$$I_d = I_{ds} + MI_{db} \tag{25}$$

Contrary to what happens with VAPDs or PMTs [136] that are ideal for single photon counting (e.g. low light spectroscopy [128]), one may well assume that the count numbers are high enough for APDs. In such case, discrete Poisson statistics may be approximated by continuous, gaussian ones. Noting that the *electron count* random variable is related to the *APD output current* random variable by means of the factor q/τ_d and that the photocurrent and dark current processes are independent, the following expressions for the mean and variance current follow:

$$\begin{cases} \mu_i = \dfrac{q}{\tau_d} M\eta\, k + (I_{ds} + MI_{db}) \\[4mm] \sigma_i^2 = \left(\dfrac{q}{\tau_d}\right)^2 M^2 F\eta\, k + \dfrac{q}{\tau_d}(I_{ds} + FM^2 I_{db}) \end{cases} \tag{26}$$

Usually, these expressions are rewritten *in terms of the current responsivity* R_i (eq.(22)), the optical input power P and over *1 Hz* bandwidth [20] noting that

$$B = \frac{1}{2\tau_d} \qquad [Hz] \tag{27}$$

If one introduces the optical power P_s, the current mean and variance become

$$\begin{cases} \mu_i = R_i P_s + (I_{ds} + MI_{db}) \qquad [A] \\[4mm] \sigma_i^2 = 2R_i^2 \dfrac{hc}{\eta\lambda} FP_s + 2q(I_{ds} + FM^2 I_{db}) \qquad \left[\dfrac{A^2}{Hz}\right] \end{cases} \tag{28}$$

Now, that both current mean and variance are have been formulated, it is worth defining the *Noise Equivalent Power (NEP)* [107]. It is defined as the incident light level impinging on a photodiode, which produces a photocurrent equal to the noise level. It is usually regarded as the most significant figure of merit for a photodetector. The NEP is a function of the photodiode's responsivity (and hence, of the wavelength λ) and the noise of the photodiode (eq.(24)). *In system applications whose SNR dominated by the photodiode performance (no thermal noise dominant), the SNR may be computed by taking the ratio of the incident optical power to the photodiode NEP.*

In general terms, using eqs.(28) and (23), NEP is defined as follows:

$$NEP(\lambda, T) = \frac{\sigma_d}{R_i} = \frac{[2q(I_{ds} + FM^2 I_{db})]^{1/2}}{\dfrac{\eta q\lambda}{hc} M} \qquad \left[\dfrac{W}{Hz^{1/2}}\right] \tag{29}$$

The nomenclature NEP(λ,T) indicates wavelength λ, and temperature, T dependence.

One step further, after multiplication by the TIA gain, G_T, the optical-to-voltage conversion can be quantified by introducing the *voltage responsivity* (see also eq.(22))

$$R_v = R_i G_T = \eta \frac{q\lambda}{hc} M G_T \tag{30}$$

Finally, inclusion of the equivalent thermal noise variance σ_{th}^2 [V²/Hz] of the preamplifier enables to express the mean and variance components of the output voltage *in terms of the voltage responsivity*. They become [127]

$$\begin{cases} \mu_v = R_v P_s + G_T(I_{ds} + M I_{db}) & [V] \\[2ex] \sigma_v^2 = 2R_v^2 \dfrac{hc}{\eta\lambda} F P_s + 2qG_T^2(I_{ds} + FM^2 I_{db}) + \sigma_{th}^2 & \left[\dfrac{V^2}{Hz}\right] \end{cases} \qquad (31)$$

But for the amplifier noise, it can be seen from eq.(31) that *the main sources of noise in the APD are* shot noises, *and originate from the primary photocurrent (noise in signal) and the dark current (rms dark noise)*. By blending eqs.(31) and (23), the signal, P_s, and dark bulk current, I_{db}, evidence a similar contribution to the shot noise variance

$$\begin{cases} \sigma_{sh,s}^2 = 2qG_T^2 FM^2 R_{io} P_s & \left[\dfrac{V^2}{Hz}\right] \\[3ex] \sigma_{sh,d}^2 = 2qG_T^2(I_{ds} + FM^2 I_{db}) & \left[\dfrac{V^2}{Hz}\right] \end{cases} \qquad (32)$$

where subscripts *sh, s, d* mean *shot, signal* and *dark,* respectively.

Two important points emerge from this equation: One is that both the power signal and the dark bulk current contribute to the shot noise exactly alike. Another is that APD noise performance is degraded by the excess noise factor, F. This parameter measures the variation in gain of both photocurrent and noise current introduced during multiplication. It can be computed using the model developed by McIntyre [148], which considers the statistical nature of the avalanche multiplication. Thus,

$$F(M) = kM + \left(2 - \frac{1}{M}\right)(1 - k) \qquad (33)$$

where k is the ionization ratio (ratio of hole to electron ionization probabilities [153]) and,

M is the APD gain factor.

It should be noted that these expressions can be used for PIN photodiodes by taking **M=1, F=1.** Fig.9 illustrates excess-noise factor behaviour vs. M.

2.2 Signal-to-noise ratio of the lidar receiver

Noise is computed in terms of the equivalent receiver system noise, B, at the TIA output. This bandwidth corresponds to the sharpest filter of the receiving chain (the antialiasing filter of the lidar receiver).

Equivalently, *SNR* can be calculated at the TIA output by defining the noise equivalent voltage density at its output, σ_{th}. It blends both the thermal noise of the TIA and that of the subsequent amplification stages (e.g. conditioning amplifiers and sampling card).
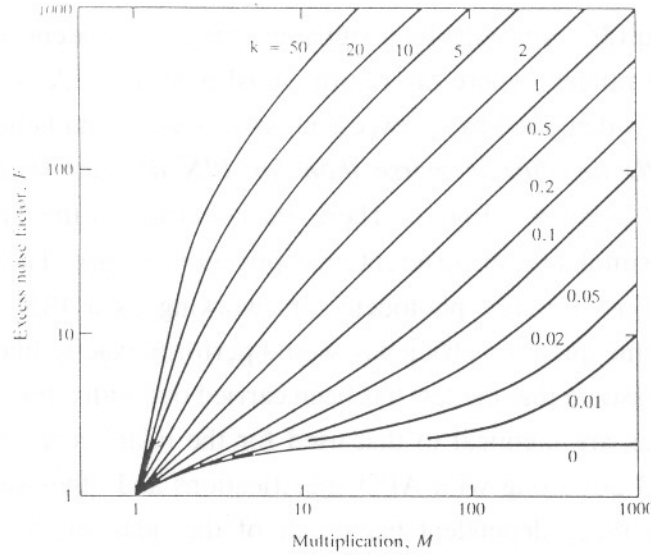
Fig. 9 Excess-noise factor calculated from eq. (33) and [20].

In the lidar receiver the signal term, P_s, blends both the range-receiver lidar signal $P(R)$ and the background power P_{back} (Chap. 2). Formally,

$$P_s = L P(R) + \xi(\lambda) P_{back} \qquad (34)$$

where $L(\lambda, R)$ represents the optical losses (Sect. 1.4), for any range R, and wavelength, λ

$$L(\lambda, R) = \xi(\lambda) \xi(R) \qquad (35)$$

The SNR can easily be computed by considering the mean voltage term in eq. (31) that contributes to the signal along with (32) and (34). It yields:

$$\frac{S}{N} = \frac{R_v L P(R)}{\left\{ 2q G_T^2 [I_{ds} + FM^2(I_{db} + R_{io}[L P(R) + \xi(\lambda) P_{back}])] B + \sigma_{th}^2 B \right\}^{\frac{1}{2}}} \qquad \frac{[V]}{[V]} \qquad (36)$$

Signal-induced shot noise, dark current shot noise and thermal noise can be rewritten as

$$\begin{cases} N_{sh,s} = \sigma_{sh,s}^2 B = 2q G_T^2 FM^2 R_{io}[L P(R) + \xi(\lambda) P_{back}] B & [V^2] \\ N_{sh,d} = \sigma_{sh,d}^2 B = 2q G_T^2 (I_{ds} + FM^2 I_{db}) B & [V^2] \\ N_{th} = \sigma_{th}^2 B & [V^2] \end{cases} \qquad (37)$$

where the subscripts *sh, th, s, d* mean *shot, thermal, signal* and *dark,* respectively.

2.3 Choosing the right detector

Obviously, one of the key parameters that has to be considered when choosing an APD for the lidar receiver is noise performance [145]. Like other detectors, an APD will be operating in one of two noise-limited detection regimes. Sensitivity (or minimum detectable power) will be limited by *dark-current shot noise or preamplifier's thermal*

noise. Regarding eq.(36), since the lidar signal is a range-dependent signal, this will always be the case at long ranges, where the return signal is weaker. As seen in Sect.2.1, APD noise performance is degraded by an excess noise factor, *F*, and hence **SNR at high power levels for and APD detector is worse than for PIN detector with the same quantum efficiency**. This is illustrated in Fig.10. There, the behaviour of the APD C30954 of EG&G Optoelectronics is simulated for different multiplication factors. The case *M=1* illustrates the behaviour of this versatile photodetector, working as a PIN photodiode. (A PIN detector with the same quantum efficiency would perform exactly like this APD with unity gain). The graphs assume that the test has been carried out with a transimpedance amplifier whose specifications are identical to that used for the lidar receiver front-end. They are listed in Tab.2 (p.3.26) along with APD specifications and other simulation parameters.

In fact, the range-dependent behaviour of the lidar signal, would condition the multiplication gain, *M*, to the exploring range. Of course, a criterion would be to somehow switch from PIN to APD mode at certain long range. Let us consider this hypothetical situation and assume a typical lidar range of *5 km*. If so, the lidar signal would last for *33 µs*. In addition, if the maximum gain is about *M=400* for a reverse polarization of V_R = *400 V*, and *M=1* is achieved at V_R = *15 V*, the high voltage source (HVS) would have to undergo a voltage swing of *385 V in only 33 µs*, every lidar shot. Obviously, this option is not possible, as seldom do HVS settling times fall below a few tenths of a second.

Evident in the curves of Fig.10 is the *knee* where SNR becomes limited by signal shot noise. Noise variances in eq.(32) (equivalently eq.(37)) are illustrated in Fig.11. Note that the intersect-point between signal-induced shot noise and thermal noise variances coincides with the *knee* in Fig.10. In the *knee* the photodiode changes from a signal shot-noise limited regime to a thermal-noise limited one.

All considered, the above controversy between choosing an APD or a PIN photodiode has been solved in favour of an APD C30954E of EG&G Optoelectronics. These are the most important pros:

1) From the beginning, a device with *acceptable quantum efficiency at both wavelengths (1064 and 532 nm)* was wished. This constraint reduced the number of possibilities very much. For this APD, typical responsivities as a function of the wavelength λ are:
   $R_i(\lambda=532nm)$ = *30.8 ± 0.5 A/W* $R_i(\lambda=1064nm)$ = *33.6 ± 0.6 A/W*
   (Figures from the device data sheets at multiplication gain *M=120*).

2) Depending on the bias voltage, the *photodetector can work either as a PIN (M=1) or as an APD (M>1)*. Therefore, a single device can cope with short and long range explorations by adjusting the optimum multiplication gain *M*.

SIMULATION PARAMETERS

λ = 532 [nm]
GT = 5750 [Ω]
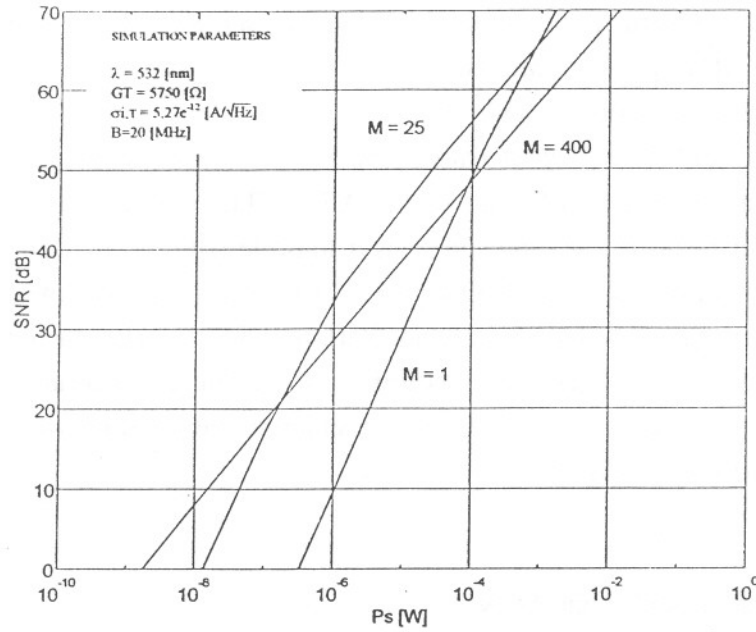$\sigma i,T$ = 5.27e^{-12} [A/$\sqrt{Hz}$]
B=20 [MHz]

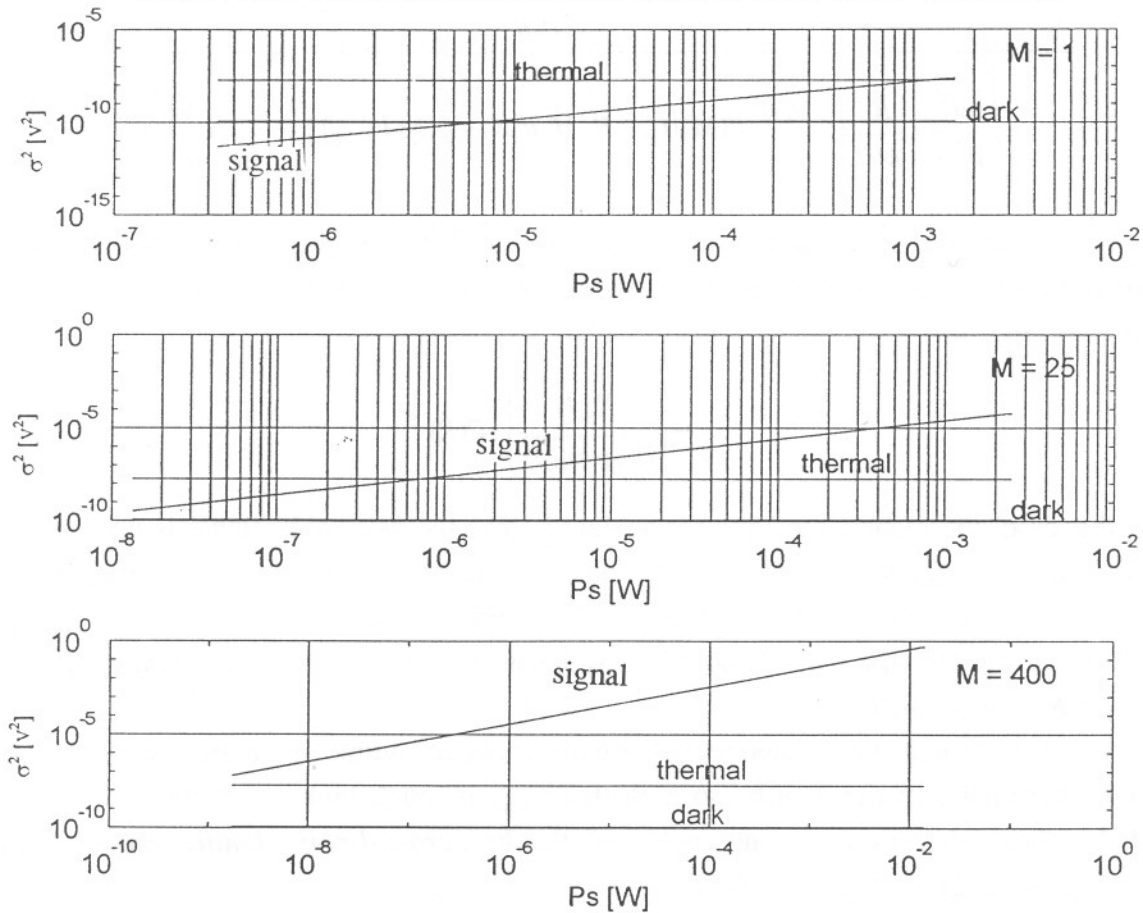*Fig.10 SNR versus optical power for different gains (APD C30954E).*

Fig.11 Noise variances for the APD EG&G C30954E.

    3) *APDs are cheaper and more robust than photomultiplier tubes (PMT)*.

    4) Such versatile photodetector represents a *sensible choice between PIN and PMT or VAPD [136] solutions*.

And this is the major con to tackle:

    1) *The field-of-view of the telescope is severely reduced* (see Tab.1) considering the small diameter (0.8 mm) of the APD. Thus, the use of a mirror to direct the laser beam to the atmosphere is a feasible solution that guarantees no overlap factor losses (eq.(16)). Otherwise, if light is output by means of a high power fiber optic coupled to some adaptation optics, overlap factor losses can be very large. This suggests the possibility of designing a second telescope with much lower focal length. That would enlarge the field-of-view of the telescope, according to eq.(13).

2.4 Noise limited modes and SNR

Dividing both numerator and denominator of eq.(36) by $R_v L$, yields

$$\frac{S}{N} = \frac{P(R)}{\left\{ \dfrac{2qI_{ds}}{R_{io}^2 L^2} \dfrac{B}{M^2} + \dfrac{2qB}{R_{io}^2 L^2} F(M)\left(I_{db} + R_{io}[LP(R) + \xi(\lambda)P_{back}] \right) + \dfrac{\sigma_{th}^2 B}{R_{io}^2 M^2 G_T^2 L^2} \right\}^{\frac{1}{2}}} \quad (38)$$

From here, four limiting situations arise [25]: The first regime or **signal shot-noise limited mode** occurs when noise due to the primary photocurrent becomes dominant. Then, eq.(38) tends to:

$$\frac{S}{N} \rightarrow \left(\frac{R_{io}L}{2qBF(M)} \right)^{\frac{1}{2}} P(R)^{\frac{1}{2}} \quad (39)$$

In this case, the highest sensitivity is achieved since the detectable signal is limited only by the quantum fluctuation of the signal itself. Note that SNR depends on the optical power as $P(R)^{1/2}$ rather than as $P(R)$ and hence SNR degrades slower. In Fig.12 this situation occurs during the first part of the lidar exploring range and cannot be guaranteed when larger R_{min} are wished.

The second noise limited regime usually arises in daytime operation. Then, the level of background radiation can be so high that P_{back} is much higher that the lidar return. Under these circumstances one speaks of the **background-noise limited mode**. Using eqs.(38) and (8) the SNR becomes:

$$\frac{S}{N} \rightarrow \left(\frac{R_{io}L\xi(R)}{2qBF(M)P_{back}} \right)^{\frac{1}{2}} P(R) = \left(\frac{R_{io}}{2qBF(M)L_b(\lambda)A_r\Omega_o} \right)^{\frac{1}{2}} \left(\frac{\xi(\lambda)}{d\lambda} \right)^{\frac{1}{2}} \xi(R)P(R) \quad (40)$$

and it is evident that the filter system function $\xi(\lambda)$ and its bandwidth $d\lambda$ play an important role in assessment of the minimum detectable power.

The third case corresponds to the *dark-current-limited mode*. Assuming a photodiode gain $M \geq 10$, the surface dark current becomes less than 1 % of its bulk counterpart. This leads to an asymptotic SNR given by

$$\frac{S}{N} \rightarrow \left(\frac{R_{io}^2 L^2}{2qBF(M)I_{db}} \right)^{\frac{1}{2}} P(R) \qquad (41)$$

Otherwise, a more complicated expression depending on both I_{ds} and I_{db} is obtained without further significative results.

Finally, the fourth case corresponds to the *thermal-noise limited mode*, the reader may feel more familiar with

$$\frac{S}{N} \rightarrow \left(\frac{R_v L}{\sigma_{th}\sqrt{B}} \right) P(R) \qquad (42)$$

This situation is clearly identified in Fig.10 at intervals where the illuminating optical power is lower. Likewise, this behaviour is little by little identified in the lidar system SNR simulation of Fig.12 as long as the range increases. For different visibility margins, all the graphs become parallel ruled by the *P(R)* dependency.

As far as the multiplication factor (M) is concerned, two different behaviour arise once these four situations are considered: *Since avalanche noise fluctuations or excess noise, F, monotonically increases with M (eq.(33)), rising M will only contribute to lower SNR in all shot-limited situations (signal-induced, background-induced, dark-induced). In a thermal-limited situation, increasing M will enhance SNR since the extra gain isolates the photodetector from the subsequent noisy stages.*

2.5 The lidar receiver and signal-to-noise ratio enhancement
2.5.1. Philosophy of the lidar receiver

For the lidar system into consideration, the goal is to maximize system range. Based on Fig.12, SNR decreases with range. At long ranges, low light levels in the return-signal justify working in the APD mode, while at short ranges, high optical levels justify the PIN mode. *Eq.(38) enables to formulate the lidar reception problem:*

Given the laser output energy, E, and the minimum exploration range, R_{min}, it is wished to:

1) find the optimum multiplication factor, M, that maximizes SNR at maximum exploration range, R_{max}, constrained to

2) the receiver's equivalent gain, G, that guarantees no saturation of the system.
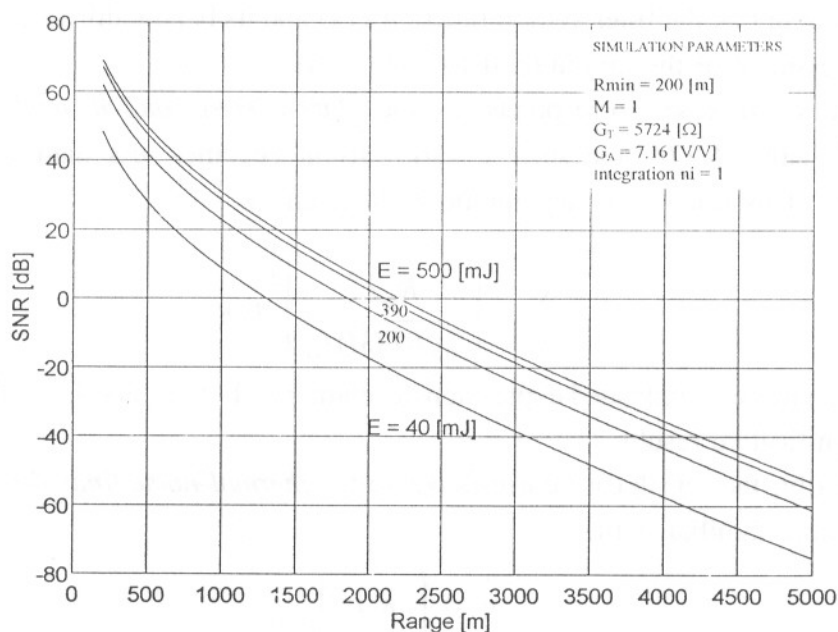
These two variables, *M* and *G*, define the *M-G domain*.

Fig.12 SNR vs. range. ($\lambda = 532$ nm, $V_M = 5$ km)

Maximum exploration range is ultimately determined by the dynamic range of the acquisition card. Since the data interval recorded will be [R_{min}, R_{max}] and SNR(R) decreases with increasing range, it is wished to optimize the SNR at the maximum range.

The equivalent gain of the lidar receiving chain should provide adequate conditioning of the lidar signal into the full dynamic range of the acquisition card. In order to do this, the return-power samples from the minimum exploration range, R_{min}, should be translated into the highest level of the acquisition card (or quite close to), while the lowest level would correspond to the absence of optical power detected.

If a $\pm V_{sat}$ bipolar card is used, the lidar receiver should ultimately perform the following conversion:

$$P(R_{min}) \to 2Vsat \qquad and \qquad no\ optical\ power \to 0$$

To achieve this final conversion, these voltage levels will have to be presented to the acquisition card with a $-V_{sat}$ offset. Even in situations where the background noise P_{back} may be neglected, samples from $R \to \infty$ will excite the lowest levels of the sampling card due to the thermal noise of the system. If one thinks in terms of equivalent noise power, this means both positive and negative spikes. So, if level translation were performed as stated, negative noise spikes would be clipped and, consequently, pulse integration would fail. This difficulty can easily be overcome by setting V_{sat} a bit lower than it should be (e.g. 5 % lower). Simulation of the AD converter with typical lidar signals, have shown that if noise is quantized with 3-4 bits, there will not be integration problems at all. This means *16 levels* of the acquisition card, and an effective dynamic range of *8 bits (256)*. Of course, it greatly depends on the SNR at range of interest. Hence, one can summarize two important specs of the lidar receiver:

a) The receiver has to be dc-coupled and must provide a variable offset with a low temperature drift. This offset should approximately be -V$_{sat}$ and adjusted during the training sequence of the lidar equipment.

b) Receiver synchronization would be desirable in order to enable the receiver at R$_{min.}$ This would avoid deep saturation of the receiver and acquisition system during the interval (0,R$_{min}$).

Fig.13 illustrates the basic block diagram of the lidar receiver. In the figure, the goal (a) is accomplished using two amplification stages: The first stage, consists on a transimpedance amplifier that performs the basic current-to-voltage conversion. The second stage is basically a conditioning amplifier. This voltage-to-voltage amplifier provides extra gain and offset to meet the dynamic range specifications of the acquisition card $(\pm V_{sat})$. In fact, this is the simplest design one can think of to meet point (a) above.

Fig.14 depicts basic noise sources in the lidar receiver under consideration.
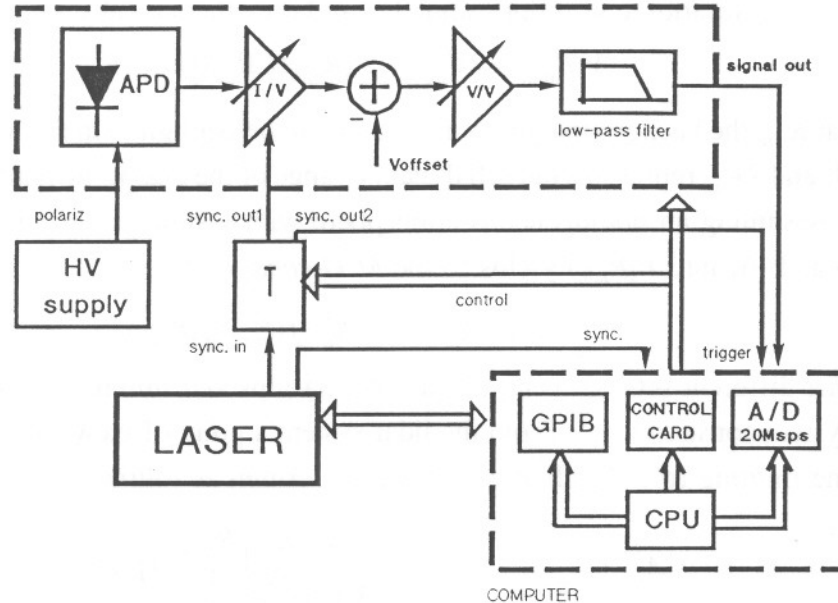
Fig.13 Block diagram of the lidar system receiver.

Now that a rough schematic of the lidar receiver has been introduced, one can return to the initial problem of the *M-G domain*. Obviously, based on Fig.13, it is wished to obtain the values (M, G_T, G_A) that fulfil the requirements (1) and (2), that is, maximum signal-to-noise ratio at R_{max} and no saturation of the system. The latter point, implies neither saturation of the acquisition card, nor saturation of any amplifier of the receiving chain.

Since these goals require a more involved formulation, the problem will be solved in two steps for clarity reasons: *In the first step, the receiving chain is assumed to be formed by the transimpedance amplifier followed by the acquisition card (Sect. 2.5.2). In the second step, the receiving chain, as sketched in Fig.13 will be discussed in Sect. 2.5.3.*
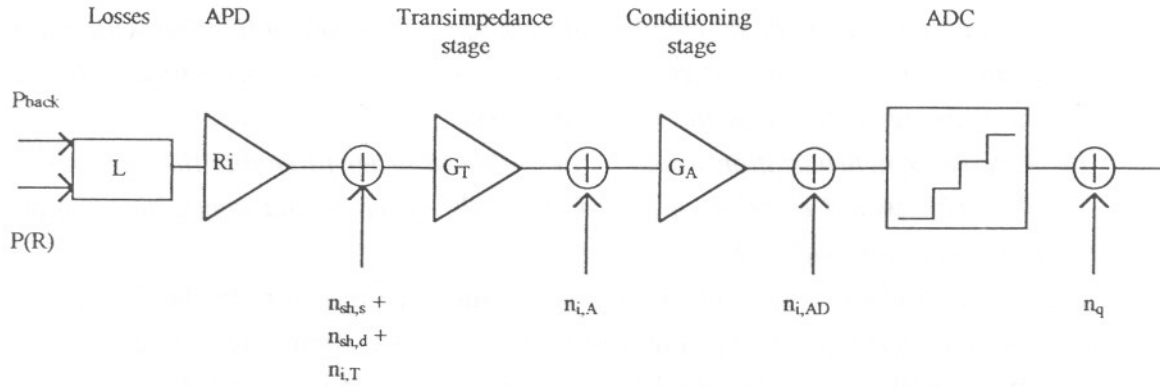
Fig.14 Lidar receiver and noise sources.

2.5.2. The M-G domain

In this section, the variable G represents the transimpedance or total chain gain.
The saturation constraint (point 1) can be mathematically stated as

$$V(R_{min}) = 2V_{sat} \rightarrow P(R_{min})LR_{io}MG = 2V_{sat} \tag{43}$$

where at R_{min} the background power is assumed to be negligible in front of the lidar return, $P(R_{min})$, and $2V_{sat}$ represents the full dynamic range of the analog-to-digital converter (AD).

Assuming an homogeneous atmosphere with visibility margin V_M, substitution of eqs.(7) and (6) into $P(R_{min})$ yields to the *M-G-constraint*

$$MG = A_o' ; \qquad A_o' = f(E, V_M, R_{min}) \tag{44}$$

Note that different sets of energy, visibility margins and minimum ranges may yield the same system constraint A_o'. From the lidar system's point of view, these three values play the same limiting role. Explicitly, A_o' can be written as follows:

$$A_o' = \frac{A_o}{E} ; \qquad A_o = \left(\frac{4V_{sat}h}{A_r L\eta q\lambda}\right)\frac{R_{min}^2}{\beta}\exp(2\alpha R_{min}) \tag{45}$$

Now, it is wished to relate the saturation constraint to the SNR given in eq.(38). Yet, thermal noise from the transimpedance amplifier, $\sigma_{i,T}^2$, and acquisition card, $\sigma_{i,AD}^2$, must be computed. For the thermal noise system variance at the transimpedance output, we have:

$$\sigma_{th}^2 = \sigma_{i,T}^2 G^2 + \sigma_{i,AD}^2 \tag{46}$$

Now, this variance can be substituted into thermal noise term of eq.(38). In doing so, one can write

$$\frac{\sigma_{th}^2 B}{R_{io}^2 M^2 G_T^2 L^2} = \frac{d}{M^2 G^2} + \frac{e}{M^2} \tag{47}$$

which is expressed in terms of the *M-G variables*.

Finally, the blending of eqs.(47) and (38) enables to write SNR as function of the multiplication gain, M and the net amplification gain G, only:

$$SNR(M,G) = \frac{P(R)}{\left\{aF(M)[LP(R)+\xi(\lambda)P_{back}]+c'F(M)+\frac{f}{M^2}+\frac{d}{M^2G^2}\right\}^{1/2}} \tag{48}$$

where:

$$a=\frac{2hcB}{\eta\lambda L^2} \quad ; \qquad b=\frac{2qBI_{ds}}{R_{io}^2L^2} \quad ; \qquad c'=\frac{2qBI_{db}}{R_{io}^2L^2}$$

$$d=\frac{\sigma_{i,AD}^2 B}{R_{io}^2L^2} \quad ; \qquad e=\frac{\sigma_{i,T}^2 B}{R_{io}^2L^2} \quad ; \qquad f=b+e \tag{49}$$

Eq.(48) represents the SNR we wanted to optimize at the beginning of this section (point 1). If this equation is linked to the saturation constraint (point 2) of eq.(44), and the following power-law approximation is used for $F(M)$ [145][123]

$$F(M) \approx M^x \quad ; \qquad 0.17 \le x \le 0.95 \tag{50}$$

the optimum gain can be obtained:

$$M = \left(\frac{2f}{ax[LP(R_{max},E)+\xi(\lambda)P_{back}]+c'x}\right)^{\frac{1}{2+x}} \quad\rightarrow\quad G=\frac{A_o'}{M} \tag{51}$$

For the photodiode considered the x-factor measured was x=0.423 (Tab.2), which is well within the interval given according to [149]. From the same reference, it is also quite close to the *x-factor* of a silicon APD *reach-through structure (x=0.3)*.

The pair M-G, written in eq.(51) represents the solution of the problem stated before. For any range R, eq.(51) dictates the best photodiode gain, that maximizes the system SNR at a given range R. In the present case, $R=R_{max}$ is the maximum range within the AD-converter dynamic range.

Fig.15 and Fig.16 illustrate the behaviour discussed. Basically, Fig.15 from [20] illustrates the optimum photodiode gain for each optical power. Graphs in Fig.16 show the M-domain according to eq.(51) that provides maximum SNR for the energies and maximum ranges shown. The plots are computed based on Tab.2 at *λ=532 nm* and *V_M=5km*. From the discussion above, energy and maximum range, R_{max}, together are equivalent to talking of a certain optical power. For this reason, the optimization could also have been done for different optical powers and Fig.15 would have been reencountered.
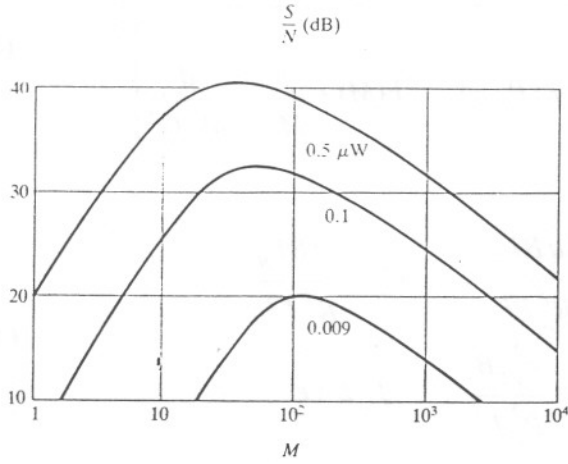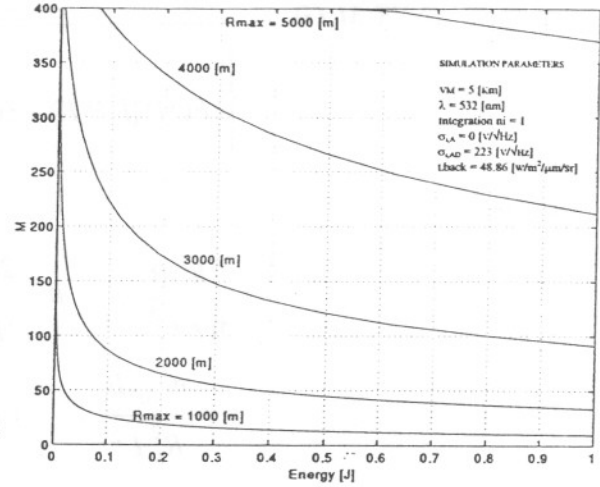
Fig.15 Sensitivity of APD detector.   Fig.16 Optimum Gain according to eq.(51).

2.5.3. Present case: the M-G2 domain

Let us consider, now, the complete lidar receiver of Fig.13. The inclusion of the conditioning amplifier following the transimpedance amplifier, must take into account these two points:

a) The receiver equivalent noise given by eq.(46) becomes

$$\sigma_{th}^2 = \sigma_{i,T}^2 G_T^2 + \sigma_{i,A}^2 + \frac{\sigma_{i,AD}^2}{G_A^2} \tag{52}$$

b) The inclusion of the conditioning amplifier may convey an additional saturation constraint if the output voltage swing of the transimpedance amplifier is smaller than the input voltage swing of the conditioning amplifier. This is to say, that under certain conditions, the acquisition card may not be saturated but the transimpedance stage can instead. *A posteriori,* the author advances that this has been the case of the lidar receiver under study. Let $V_{sat}=\pm 1$ be the maximum output swing of the transimpedance stage. Then, the *amplifier saturation constraint* can be written as

$$P(R_{min})LR_{io}MG_T \le V_{sat,1} \quad \rightarrow \quad MG_T \le A_1 \tag{53}$$

As in Sect.2.3.2, the goal is to maximize the SNR

$$SNR(M,G_T,G_A) = \frac{P(R)}{\left\{aF(M)[LP(R)+\xi(\lambda)P_{back}]+c'F(M)+\dfrac{b_1}{M^2}-\dfrac{d_2}{M^2 G_T^2}+\dfrac{d_3}{M^2 G_T^2 G_A^2}\right\}^{1/2}} \tag{54}$$

or, equivalently, to minimize the noise variance. Substitution of (52) into (38) yields the *formulation of the minimization problem*

$$\min(\sigma^2)\big|_{M,G_T,G_A} = \left\{ a_1(R)M^x + \frac{b_1}{M^2} + \frac{d_2}{M^2 G_T^2} + \frac{d_3}{M^2 G_T^2 G_A^2} \right\} \tag{55}$$

in the feasible domain given by the constraints:

    - *eqs.(44), (53) or saturation constraints* and,

    - $M \in (1, M_{max})$; $G_T \in (t_{min}, t_{max})$; $G_A \in (a_{min}, a_{max})$ or *gain constraints.*

where

$$d_1 = \frac{\sigma_{i,T}^2}{R_{io}^2 L^2} \;\; ; \qquad d_2 = \frac{\sigma_{i,A}^2 B}{R_{io}^2 L^2} \;\; ; \qquad d_3 = \frac{\sigma_{i,AD}^2 B}{R_{io}^2 L^2} \;\; ; \tag{56}$$

$$a_1 = a[LP(R) + \xi(\lambda)P_{back}] + c' \;\; ; \qquad b_1 = b + d_1$$

The constants a, b, c' have already been defined in eq.(49). The variable G of eq.(44) represents net amplification gain due to both transimpedance and conditioning stages, $G = G_T \cdot G_A$.

The gain constraints depend on the minimum and maximum gains of the amplification stages. Note that eq.(55) is the counterpart of the denominator of eq.(48).

    The tri-dimensional constrained optimization is solved by link-budget software whose code is MATLAB based.

2.6 Default values concerning this section

    Tab.2 summarizes default values of the simulations presented.

NOTES: 1) Link-budget simulations have considered worst case noise variances. They assume maximum gains in the amplifiers and hence simulations offer conservative lidar range estimations.

        2) Link-budget simulations shown here assume: $\eta = 54.9\ \%$, $R_{io} = 237$ mA/W and $\lambda = 532$nm.
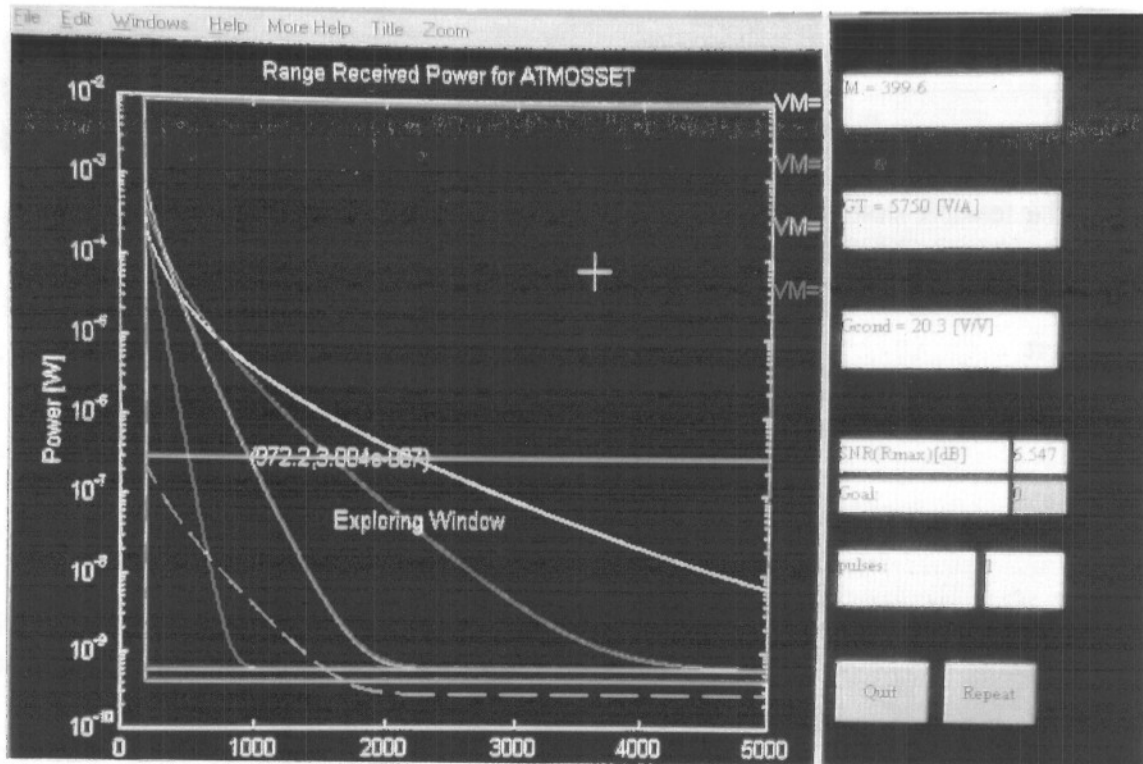
Fig.17 Exploring window and SNR optimization in link-budget software.

PHOTODETECTOR PARAMETERS

APD parameters (measured values unless otherwise indicated)

Type: C30954E EG&G Optoelectronics. Silicon, double-diffused *reach-through* structure.

Diameter: ⌀0.8 mm   (specs)

Quantum efficiency: $\eta = 54.9 \pm 2.3$ % ($\lambda_1 = 532$nm) $\eta = 30.0 \pm 1.3$ % ($\lambda_2 = 1064$nm)

Intrinsic responsivity (eq.(23)): $R_{io} = 237 \pm 12$ mA/W (λ_1) $R_{io} = 258 \pm 13$ mA/W (λ_2)

Gain performance: a) multip. factor (eq.(22)): $M \epsilon [1$ to $400]$ (for PINs, $M=1$).

                        b) k-factor (eq.(33)): $k = 0.033$

                        c) x-factor (eq.(50)): $x = 0.423$

Dark current(eq.(25)): a) surface dark current: $I_{ds} = 13.8$ nA

                        b) bulk dark current: $I_{db} = 6.0$ pA

Noise density (eqs.(24),(27)): $\sigma_d \approx 2$ pA·Hz$^{-1/2}$, $M=400$ (specs: $\sigma_d \approx 1$ pA·Hz$^{-1/2}$, $M=120$)

Noise Equivalent Power (eq.(29)):   **NEP$_{532}$ = 21 fW/Hz$^{1/2}$** ($M=400$) $B \approx 500$ MHz

Transimpedance amplifier specs (measured values unless otherwise indicated)

Gain: Variable from 750 to 5750 Ω

Equivalent input noise: $\sigma_{th} = 5.27$ pA·Hz$^{-1/2}$ (at maximum gain, $G_T = 5750\Omega$)

Equivalent noise bandwidth: $B = 20$ MHz

Tab.2 Defaults used link-budget simulations of Sect.2.

3. LIDAR RANGE ESTIMATION

3.1 Absolute minimum and maximum exploring ranges of the system

Eqs.(44) and (53) or saturation constraints enable to mathematically formulate the absolute minimum system range, $R_{min}{}^a$. The superscript a indicates absolute. *For each output energy, E, $R_{min}{}^a$ is determined by saturation of either an amplification stage or the acquisition card itself when all the system gains (M, G_T, G_A) are set to minimum.* Obviously, which of the two constraints becomes dominant depends on the input and output swings of the components involved in the receiving stages. Yet, let us assume a ± 1-*V* input range for the acquisition card, a minimum conditioning gain of less than *3V/V* and a ± 3-*V* output swing for the transimpedance amplifier. These ranges are pessimistic considering the *state-of-the-art* in operational amplifiers [135][134][106]. In the example, it turns out that only acquisition card constraint of eq.(44) applies. Note that $R_{min}{}^a$ attains the lowest value when the system gain is minimum. Fig.18 plots absolute minimum system ranges for different output energies and visibility margins based on Tab.3.

On the other hand, *the absolute maximum system range, $R_{max}{}^a$ is determined by the locus of signal-to-noise ratio equal to some lower bound, SNR.* This lower bound is usually *SNR=1,* even though signal-to-noise ratio at $R_{max}{}^a$ can be improved by pulse integration. Since the return signal from $R_{max}{}^a$ is very dim, there is no significative signal-induced shot noise (eq.(37)). Under these conditions, *SNR is thermal-noise limited* (eq.(42)). Then, $R_{max}{}^a$ can be computed by setting $SNR(R_{max}{}^a)=1$ in eq.(54) or equivalently, from the system NEP_s. This concept is similar to the photodiode NEP given in eq.(29). Reconsidering eq.(37), NEP_s can be written as

$$NEP_s = \frac{\left[(\sigma_{sh,d}^2 + \sigma_{i,T}^2)\,G_T^2 G_A^2 + \sigma_{i,A}^2 G_A^2 + \sigma_{i,AD}^2\right]^{1/2}}{R_{io} M G_T G_A} \qquad \left[\frac{W}{\sqrt{Hz}}\right] \qquad (57)$$

Note that NEP_s is minimum at maximum system gain. It is important to emphasize that this affirmation does not contradicts Sect.2.5. There, it was suggested that for any $R_{max} < R_{max}{}^a$ one can enhance SNR by adjusting the right multiplication gain, M. In the optimization procedure, multiplication gains below the maximum gain suggest that the system is still far from the thermal-noise limited mode. Otherwise, if the optimization were performed at $R_{max}{}^a$ range, which is very often of several kilometers, the result would have been that all gains should be as high as possible in order to counteract thermal noise. Quite often, this well-known result from the Friis formula is taken for granted [7].

Similarly to Fig.18, Fig.19 plots absolute maximum range versus output energy and visibility margin. Fig.20 represents the differential range $R_{max}{}^a$-$R_{min}{}^a$.

It is important to warn here, that $R_{min}{}^a$ and $R_{max}{}^a$ do not take into account optical limitations, mainly overlap factor (Sect.1.2). For the time being, this factor is assumed unity over the entire range ($R_{min}{}^a$,$R_{max}{}^a$). Otherwise, the received signal will be distorted.
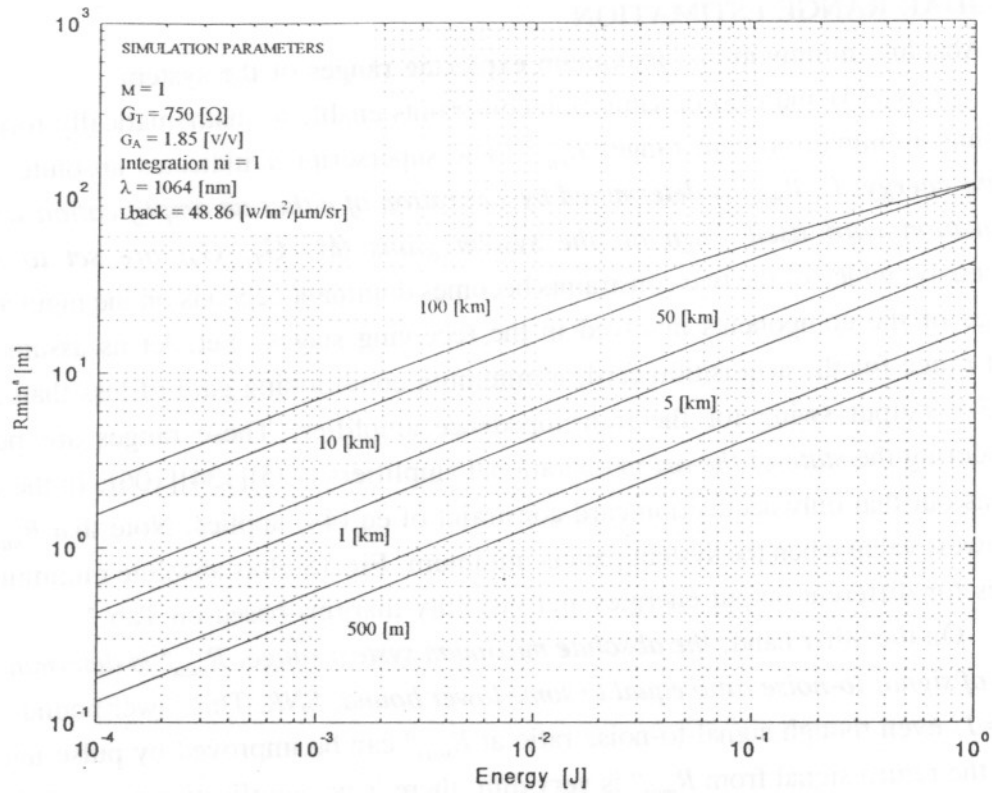
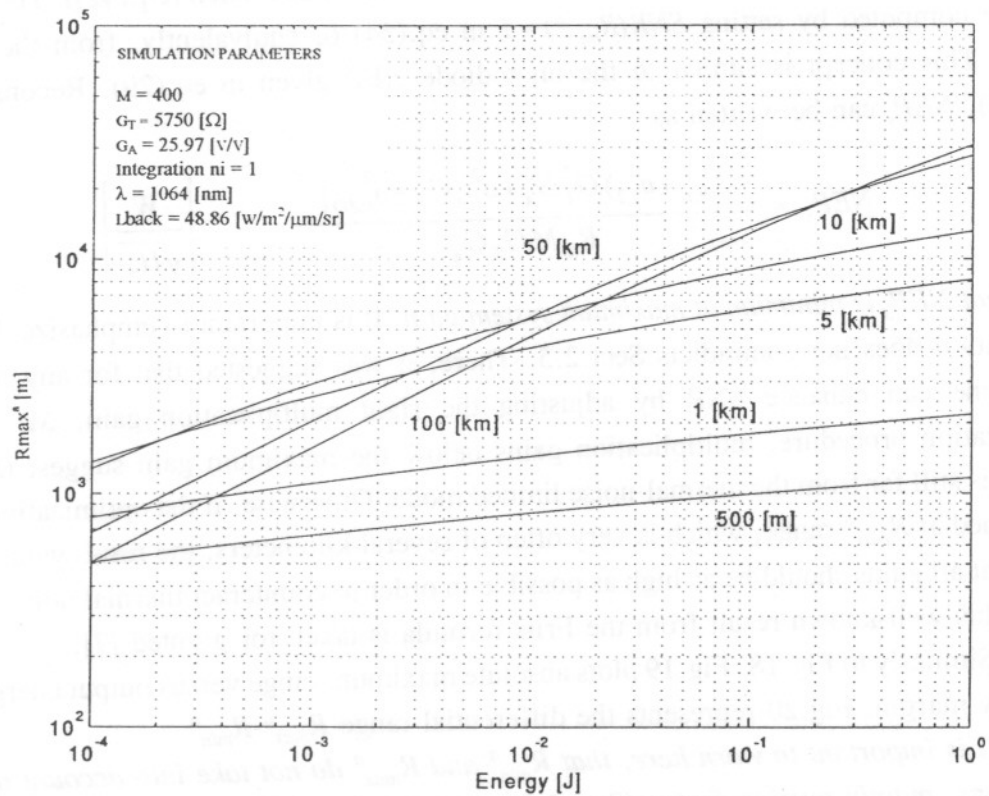*Fig.18 R_{min}^{a} vs. energy and visibility margin.*

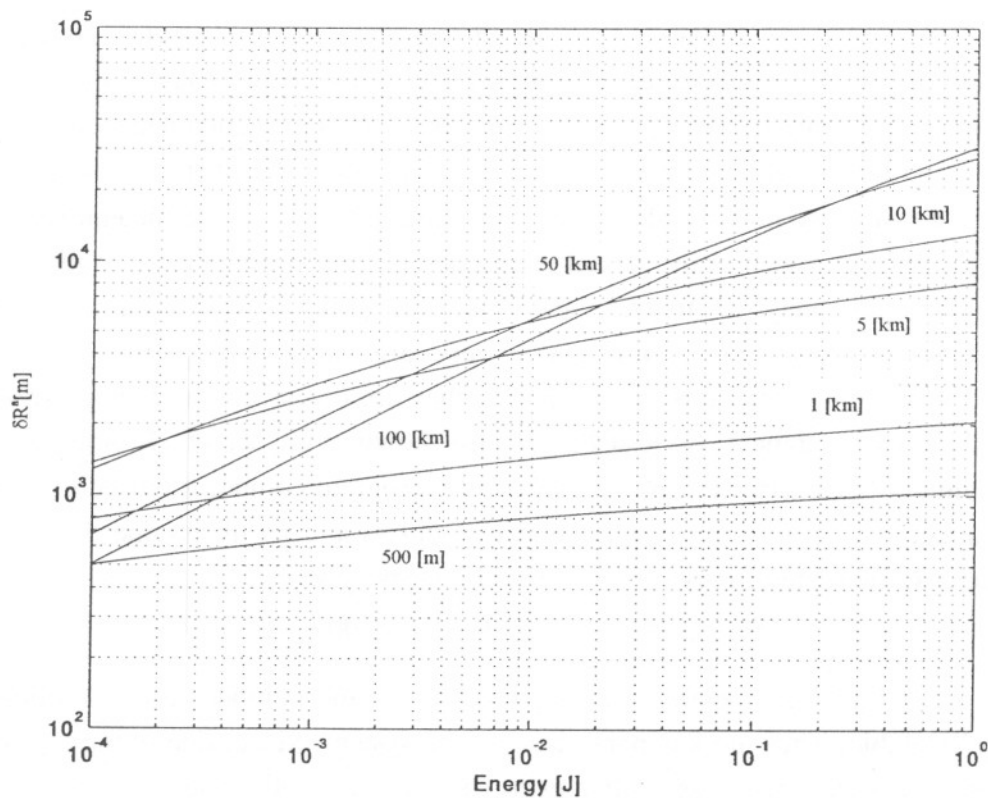*Fig.19 R_{max}^{a} vs. energy and visibility margin.*

Fig.20 Differential range $\Delta R^a = R_{max}{}^a - R_{min}{}^a$ *vs. energy and visibility margin.*

3.2 The exploring window

Absolute system ranges are sometimes unpractical limits. In fact, they assess maximum system performance given some knowledge about the visibility margin. The visibility turns out to be an additional uncertainty factor unless transmittance measurements are readily available. *In addition, the exploration cannot directly be made from $R_{min}{}^a$ to $R_{max}{}^a$ because of the limited dynamic range of the acquisition system. One can see the acquisition system as window that explores a determinate segment (from R_{min} to R_{max}) of the absolute system ranges, $R_{min}{}^a$ to $R_{max}{}^a$. The synchronization system anchors the window to the desired R_{min}.* The larger the dynamic range, the larger the segment. Since every bit of and AD-converter equals 6 dB dynamic range, a card with 12 bits will have and input range of 72 dB. This outweighs the specifications of even the best logarithmic amplifiers in the market [115]. Since the logarithmic conversion can ultimately be made in a computer, this option has been discarded in favour of a card with 12 or more bits. Tab.3 lists some specifications of the card chosen.

Fig.17 shows many features of the exploring window in *link-budget* software. Basic inputs to the exploring window procedure are laser energy and a set of visibility margins, the user will choose from. Other system parameters are taken from system default files. Maximum and minimum system detectable powers are indicated by top and bottom lines. These values are absolute limits and correspond to minimum and maximum system gains, respectively.

The exploring window is represented there by a rectangle, whose top left corner the user can anchor at the wished R_{min} and V_M curve, using the mouse. Once the user has chosen these two parameters, R_{max} is computed from the intersection between the exploring window and the power curve for that visibility, V_M. After that, the system gains, M, G_T, G_A are computed according to the M-G2 procedure (Sect.2.5.3) and the equivalent system noise (eq.(37)) is drawn in the screen with a dashed line. Note that this noise, especially at short ranges, may be much higher than the power levels obtained from NEP_s computations because shot noise is usually very large at these ranges.

Once the system gains are computed, they are displayed and they are ready to be set in the lidar equipment. SNR at R_{max} is also shown in the screen. From this value and the minimum SNR wished at R_{max}, SNR_{goal}, the program computes the number of pulses to be integrated, if any. Let n_i the number of pulses and let us assume that the SNR is measured at the system output as a ratio of voltages. Then,

$$10\log(n_i) = SNR_{goal} - SNR(R_{max}) \tag{58}$$

After that, the system gains, M, G_T, G_A and the number of integrations, n_i, are readily available. They are key parameters of the system. Since, due to the uncertainty in the visibility margin the sought-after gains can change, remote digital control over them would be highly valuable. For example, a GPIB high voltage power supply could serve the purpose to control the APD gain and a programmable lidar receiver should be considered. Of course, the lidar system makes do with analog control over its amplifiers (e.g. multi-turn potentiometers) but doing so, would make the *M-G2* optimization cumbersome, if not impossible.

All in all, the tricky parameter in the optimization procedure is the visibility margin. Ideally, it has to be determined from in-situ measurements, e.g. transmitometers or balloon-borne measurements. Otherwise, work with a digitally controlled lidar equipment will provide valuable expertise.

3.3 Gain and synchronism specs

Maximum detectable level complies with the acquisition card constraint of eq.(43). Yet, minimum detectable power may well be buried into the system noise level. If so, thermal noise realizations will have to be integrated in order to retrieve the signal. Moreover, if the integration is wished to succeed, noise clipping must be avoided, otherwise noise statistics will no longer be gaussian. Based on simulations, if minimum detectable power excites $\kappa > 8$ levels of the digitizer, clipping is virtually eliminated. This factor κ is called the *crest-factor*. For the maximum and minimum power, one can write:

$$P_{max} = \frac{2V_{sat}}{LR_{io}M_{min}G_{min}} \qquad P_{min} = \left(\frac{2Vsat}{2^{bits}-1}\kappa\right)\frac{1}{LR_{io}M_{max}G_{max}} \qquad (59)$$

For the *GaGe Compuscope 1012 digitizer* (Tab.3), V_{sat} can be programmed within the range *5 V, 1 V, 500 mV, 200 mV and 100 mV*. From the beginning, it was agreed that the design should be carried out using operational amplifiers. Considering that typical output voltage ranges for them are between ± 6, ± 13 V *(no load)* and between ± 4.5, ± 8 V *(100-Ω load)* [134][135], the saturation voltage of the card was chosen to be $V_{sat} = 1V$. This allows conventional system design based on typical 75 or 50Ω impedances.

Next, let us give some insight on how the receiver gain can be assessed: In Fig.7 (E=500 mJ, λ=532 nm), P_{min} is less than *1 mW* for $R_{min}{}^a \approx$ *200 m*. This minimum range may be considered a sensible choice since lower ranges may well fall into the range interval where the overlap factor has not settled yet (case 3 in Sect.1.2). Anyhow, let us consider that P_{min} may be as high as *5 mW* under a more conservative criterion.

As for $R_{max}{}^a$, the design specification is $R_{max}{}^a \approx$ *5 km* under average seeing conditions. This goal has been set because $R_{max}{}^a$ is well above the boundary layer and covers most of the meteorological phenomena of interest.

Since the higher range is the most difficult range to assess *a priori*, during the link-budget stage of the project a conservative visibility margin was considered. Even though V_M lies between *15 and 23.5 km* (Fig.6) for average seeing conditions, $R_{max}{}^a$ condition was computed with a V_M as low as *5-10 km (light-moderate haze), given an output energy of 1 J* (λ=1064 nm and Fig.8). Under these assumptions and the plots of Fig.7, P_{min} is about *1-5 nW* for E = 500 mJ and about *2-10 nW* for E = 1J, since the APD has similar responsivities at both operating wavelengths. Looking to these figures, P_{min} = *5 nW*, seems a reasonable choice. Summing up, the system is specified to cope with approximate power levels in the interval:

$$P \in (5 \ mW, \ 5 \ nW).$$

Assuming *L = 0.3, R_{io}=0.3 (specs), M$\in$(1,120) (specs), V_{sat} = 1V, κ=8*, eq.(59) yields the following MG and G gains, where $G=G_TG_A$:

$$MG \leq 4400 \ \Omega; \ MG \geq 8.7 \times 10^6 \ \Omega$$
$$G_{min} \leq 4400 \ \Omega; \ G_{max} \geq 72 \times 10^3 \ \Omega$$

These results are also shown in Tab.3. Note that the final performance of the system will ultimately depend on the measured multiplication gain, M. In fact, Tab.2 and Tab.3 evidence a mismatch between the maximum multiplication gain reported by the manufacturer and the measured one.

To have an estimate of the system NEP, NEP_s, G_{max} can be split as G_T= *4000 Ω* and G_A= *18 V/V*. If the values specified (Tab.2,Tab.3) are considered *M=120, $\sigma_{i,T}$=5.27 pA·Hz$^{-1/2}$, $\sigma_{i,A}$=3.41 nV·Hz$^{-1/2}$, $\sigma_{i,AD}$=90 nV·Hz$^{-1/2}$*, the system NEP becomes, NEP_s=220 fW·Hz$^{-1/2}$, which under a noise equivalent bandwidth of 9.2 MHz, leads to a noise power

of 0.6 nW (SNR= 9.2 dB). Note that the same $R_{max}{}^a$ could have been derived from a family of plots similar to Fig.19 using the criterion $SNR(R_{max}) = 9.2\ dB$. If the measured multiplication gain, M=400, is considered (Tab.2,Tab.3) $NEP_s = 70\ fW \cdot Hz^{-1/2}$.

As to the synchronization specs, R_{min} is well beyond our purposes taking $R_{min}{}^s = 100$ m and $R_{max}{}^s = 2\ km$. *The superscript s* refers to synchronism. Obviously, these limits are overspecified but there are reasons for that: First, $R_{min}{}^s$ must be lower than $R_{min}{}^a$ so as to leave the overlap factor as the only limiting factor at R_{min}. In addition, it is important to have $R_{min}{}^s$ as low as possible to help time-range calibration of the system using practical fibre optic lengths available in the laboratory. The lower limit of $R_{min}{}^s$ is caused by intrinsic delays of the synchronization unit that can no longer be reduced (e.g. propagation delays). As for $R_{max}{}^s$, exploration can begin at large ranges, such as *2 km,* and set the system gain to cope with *far range scans*. At long ranges signal-to-noise ratios may become negative, but it is precisely at these ranges where pulse integration is welcomed.

All in all, $R_{min}{}^s$ and $R_{max}{}^s$ are equivalent to time-delays of 0.66 μs and *13.3 μs,* respectively.

RECEIVER AND SYNCHRONISM PARAMETERS

Acquisition card: GaGe CompuScope 1012, 20 Msps, 12 bits
    Dynamic input range: 72 dB (6 dB/bit)
    Input scales: ±5V, **±1V** (range chosen), ±500 mV, ±200 mV, ±100 mV
    Noise: $\sigma_{i,AD} = 223\ nV \cdot Hz^{-1/2}$ (measured), 90 nV·Hz$^{-1/2}$ (specs[*])

Receiver: (link-budget computations assume visibility margin, $V_M \approx 5\ km$)
    Estimated system absolute ranges and typical power levels (Fig.7, Fig.8):
        $R_{min}{}^a < 200\ m$ (P $\approx$ 5 mW) $R_{max}{}^a \geq 5\ km$ (P $\approx$ 5 nW)
  Assuming $L=0.3$, $R_{io}(specs^*)=0.3\ A/W$, $M(specs^*)\epsilon(1,120)$, $V_{sat}=1V$, $\kappa=8$ (eq.(59))
    Chain gain, M·G (G=G_T·G_A):   MG $\leq$ 4400 Ω; MG $\geq$ 8.7·10^6 Ω
      if M ϵ (1,120) then $G_{min} \leq$ 4400 Ω; $G_{max} \geq$ 72·10^3 Ω
    Further specs: Variable gain between G_{min}, G_{max} and offset (-1 V level shift).
    Noise: a) transimpedance stage $\sigma_{i,T} = 5.3\ pA \cdot Hz^{-1/2}$ (measured)
          b) conditioning stage $\sigma_{i,A} = 3.4\ nV \cdot Hz^{-1/2}$ (measured)
    $NEP_s \approx$ 220 fW·Hz$^{-1/2}$ (eq.(57))   at (MG)$_{max}$ (M=120, G_T=4000Ω, G_A=18 V/V)
    $NEP_s \approx$ **70 fW·Hz$^{-1/2}$** (measured)   (as above but using measured M=400)

Synchronization unit:
    $R_{min}{}^s$ = 100 m (0.66 μs) $R_{max}{}^s$ = 2 km (13.3 μs)

[*] The note *(specs)* indicates that the value has been taken from the device data sheet. Measurement of the parameter at a later time revealed a different figure. The new figure is indicated as *(measured)* into brackets. Most figures have been advanced from subsequent chapters. The plots presented in this chapter include updated values and show actual performance.

Tab.3 Defaults used in link-budget simulations of Sect.3.

4. A glance to link-budget software

Link-budget software, from which all the plots presented in this chapter have been processed, offers many possibilities to study lidar systems. It basically concentrates on the optical and electric stages of the lidar receiver. *Link-budget* along with *link-atmos* (radiance and atmospheric studies) and *link-detect* (inversion algorithms) build the *link-lidar* main package developed during this PhD thesis. About link-budget possibilities:

1) *Defaults*. Lidar system parameters and photodiode libraries can be readily up-to-dated using pop-up menus. For examples, visibility pop-up menus include tabulated values of (V_M, α, β) vs. wavelength.

2) *Sweep range*. Absolute system ranges, R_{min}^{a}, R_{max}^{a} are assessed. Results can be parameterized for any lidar parameter and hence effects on the overall system performance can easily be pinpointed.

3) *Lidar and the atmosphere*. Exploring window (Sect.3.2) and signal-to-noise studies are included.

4) *Overlap factor*. (Sect.1.2).

5) *Stochastic simulator*. Electrical signals at different points of the lidar receiver of Fig.13 are simulated. Random noise generators simulate APD and amplifier noises and hence SNR can be estimated.
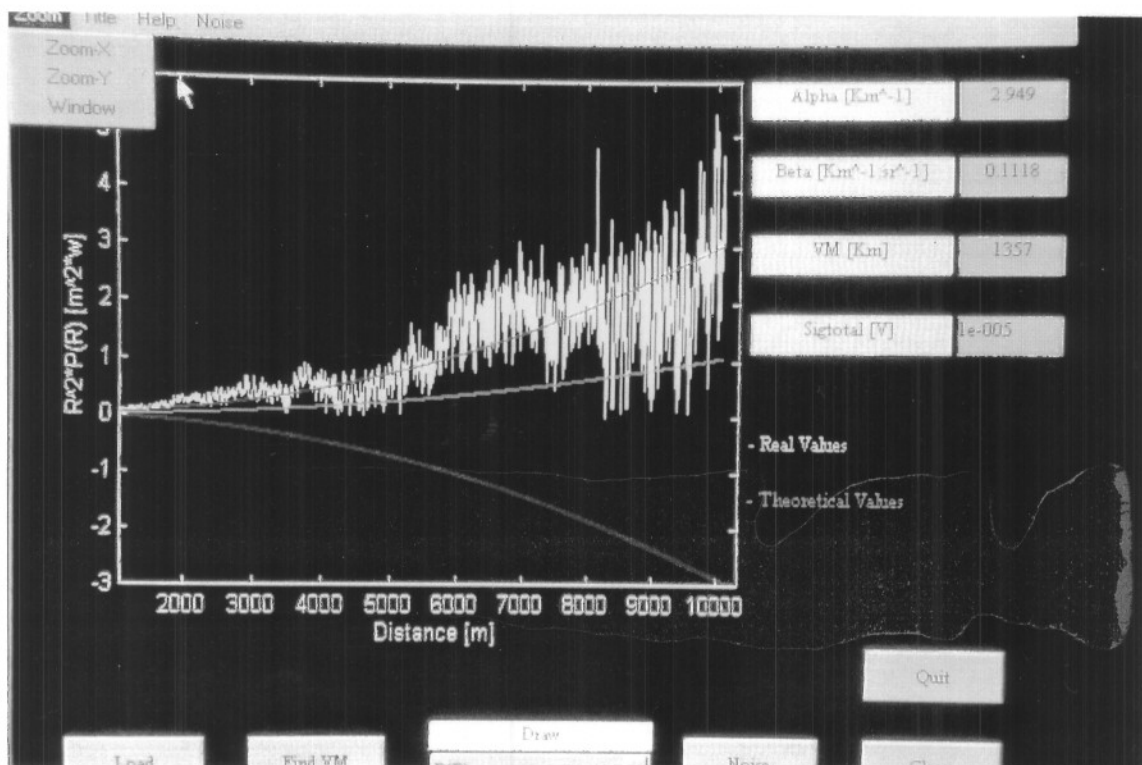

Fig.21 Link-budget software used in lidar design.